Information Sciences 231 (2013) 4–14

Contents lists available at SciVerse ScienceDirect

Information Sciences

journal homepage: www.elsevier.com/locate/ins

Toward a more practical unsupervised anomaly detection system

Jungsuk Song [a,d,*], Hiroki Takakura [b], Yasuo Okabe [c], Koji Nakao [a]

[a] National Institute of Information and Communications Technology, 4-2-1, Nukui-Kitamachi, Koganei-shi, Tokyo 184-8795, Japan
[b] Information Technology Center, Nagoya University, Furo-cho, Chikusa-ku, Nagoya 464-8601, Japan
[c] Academic Center for Computing and Media Studies, Kyoto University, Sakyo-ku, Kyoto 606-8501, Japan
[d] Korea Institute of Science and Technology Information, 245 Daehangno, Yuseong, Daejeon, 305-806, Korea

ARTICLE INFO

Article history:
Available online 22 August 2011

Keywords:
Intrusion Detection System
Clustering
One-class SVM
Anomaly detection

ABSTRACT

During the last decade, various machine learning and data mining techniques have been applied to Intrusion Detection Systems (IDSs) which have played an important role in defending critical computer systems and networks from cyber attacks. Unsupervised anomaly detection techniques have received a particularly great amount of attention because they enable construction of intrusion detection models without using labeled training data (*i.e.*, with instances preclassified as being or not being an attack) in an automated manner and offer intrinsic ability to detect unknown attacks; *i.e.*, 0-day attacks. Despite the advantages, it is still not easy to deploy them into a real network environment because they require several parameters during their building process, and thus IDS operators and managers suffer from tuning and optimizing the required parameters based on changes of their network characteristics. In this paper, we propose a new anomaly detection method by which we can automatically tune and optimize the values of parameters without predefining them. We evaluated the proposed method over real traffic data obtained from Kyoto University honeypots. The experimental results show that the performance of the proposed method is superior to that of the previous one.

1. Introduction

With the rapid development of internet infrastructure and local networks, security threats such as distributed denial of service (DDoS), computer viruses and internet worms, are simultaneously growing for computer systems and networks. Many efforts have been taken to counteract these cyber attacks in the last decade; including cryptography, firewalls and Intrusion Detection Systems (IDSs). Among these techniques, IDS [6,2] is becoming increasingly significant in maintaining proper network security and defending crucial computer systems and networks from malicious attacks [1].

There are mainly two types of Intrusion Detection Methods: misuse detection and anomaly detection. In misuse detection-based IDSs, network traffic is monitored to detect illegal access using predefined attack signatures [13,10]. If it is observed that any intrusion or suspicious activities match patterns of predefined signatures, they raise the corresponding alerts. On the other hand, IDSs based on the anomaly detection method use normal patterns to detect abnormal activities from observed data. They typically attempt to identify deviations from predefined normal patterns, and regard them as potential attacks. The former can detect well-known attacks at relatively higher accuracy than the latter, but they have a fatal weakness in that they cannot detect unknown attacks (*i.e.*, 0-day attacks) that are not matched to any predefined signatures. Anomaly detection-based IDSs can potentially detect unforeseen attacks since their activities differ from normal patterns.

* Corresponding author. Address: National Institute of Information and Communications Technology, 4-2-1, Nukui-Kitamachi, Koganei-shi, Tokyo 184-8795, Japan. Tel.: +82 42 869 0729; fax: +82 42 869 1119.
E-mail addresses: song@kisti.re.kr (J. Song), takakura@itc.nagoya-u.ac.jp (H. Takakura), okabe@i.kyoto-u.ac.jp (Y. Okabe), ko-nakao@nict.go.jp (K. Nakao).

0020-0255/$ - see front matter © 2011 Elsevier Inc. All rights reserved.
doi:10.1016/j.ins.2011.08.011

During the last decade, many machine learning and data mining techniques have been applied to IDSs to improve their performance and construct them with low cost and effort. In particular, *unsupervised* anomaly detection techniques [18,7,8,16,15,22,20,31] have received a great deal of attention because they can construct intrusion detection models without using labeled training data (*i.e.*, with instances preclassified as being or not being an attack) in an automated manner. With their intrinsic ability to detect 0-day attacks, this advantage makes them more easily applied to real environments since labeled data or purely normal data cannot be readily obtained in practice.

Instead of using unlabeled training data, they require several parameters (*e.g.*, the ratio of attack data to normal data and a threshold to determine attack data and normal data) during their building process. Since normal patterns on distinct networks differ from each other, this drawback leads to efficiency deterioration when we deploy them into diverse network environments. In other words, IDS operators and managers must carefully tune and optimize parameters required for building intrusion detection models based on changes of their network characteristics. Considering the generality of misuse detection-based IDSs where attack signatures can be used for any IDSs operating on distinct networks, we need to cope with this fatal weakness in unsupervised anomaly detection techniques.

In previous research [20], we have proposed an unsupervised anomaly detection technique that outperforms those in existing research. However, it needs three parameters (α, β and k) for constructing an intrusion detection model as described in Section 2. Among the three parameters, we focus on only two α and k, because we have proven that parameter β barely affects the performance of the previous anomaly detection method as long as we choose a reasonable range of value in [20]. In this paper, we propose a new anomaly detection method by which we can automatically obtain the optimize values of two parameters without predefining them. We evaluated the proposed method over a real dataset of network connections gathered from Kyoto University honeypots [21]. The experimental results show that the proposed method is superior to the previous one.

The rest of the paper is organized as follows. Section 2 gives a brief description of the previous anomaly detection method. Section 3 describes the proposed anomaly detection model in detail, and Section 4 gives the experimental results including their analysis. Section 5 presents concluding remarks and suggestions for future study.

2. Previous anomaly detection method

Fig. 1 shows the training and testing phases of the previous anomaly detection method. The training phase is composed of three main steps: filtering, clustering and modeling. The training phase first filters out most of the attack data from the original training data (①Filtering) to guarantee that the majority of each cluster obtained from the following clustering process consists of normal data, and partitions the filtered training data into k clusters which indicate normal patterns (②Clustering). Clustering is carried out due to the existence of different types of normal patterns in the traffic data such as HTTP, SMTP and FTP. Thus, it is very important to cluster the traffic data into individual clusters to more accurately identify their activities. One-class SVM [16,4] is then applied to each normal cluster (③Modeling), and k SVM models (*i.e.*, k hyperspheres) are

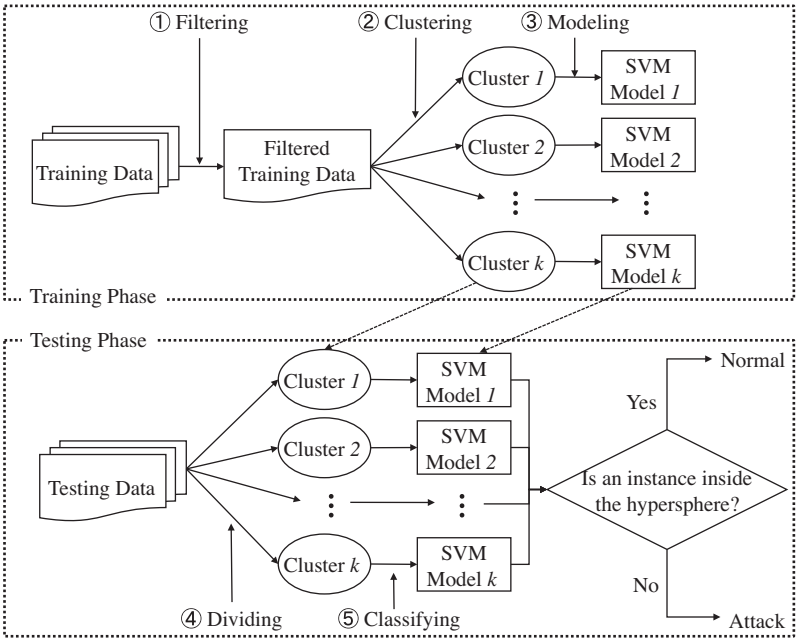

Fig. 1. Training and testing phases.